Overview of the Risk-aware Machine Learning Project

FY17 President's and Director's Fund Project

Hiro Ono
Jet Propulsion Laboratory, California Institute of Technology

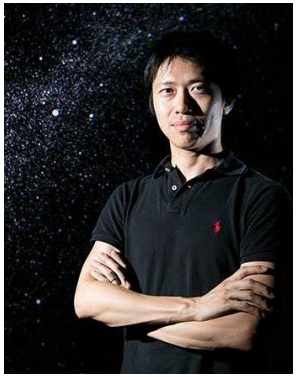
The Team

Caltech PI: Prof. Yisong Yue

Dr. Yanan Sui

Jialin Song

Hoang Le



JPL PI: Dr. Hiro Ono

Ravi Lanka Subrahmanya

Dr. Mitch Ingham

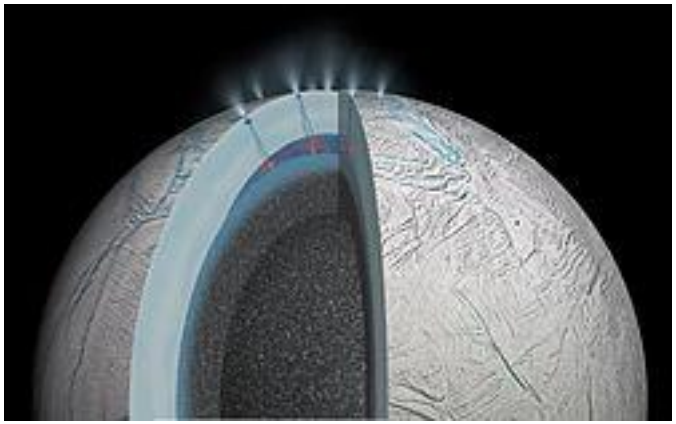
Dr. Tara Estlin

External collaborators:

Akifumi Wachi (University of Tokyo)

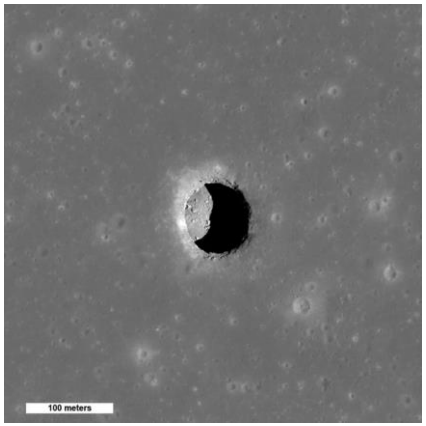
Tiago Vaquero (MIT)

Why Risk-Aware Autonomy?



Icy moon's subsurface ocean

Surface of Venus



Lunar/Martian caves

- Frontiers of Solar system exploration are characterized by:
  - Limited/no orbital reconnaissance
  - Very limited communication
  - Limited lifetime of spacecraft (e.g., Venus: few hours)
  - Unknowns and risks
- Spacecraft must:
  - Identify & assess risks on-board
  - Autonomously choose actions to balance risk & science gain
- Related talks:
  - Automated underwater vehicles: Eric Timmons (MIT), 1:20pm
  - Small satellites: Akifumi Wachi (U Tokyo), 1:40pm
  - Mars rover: Raymond Francis (JPL), 2:50pm

What is Risk-aware Autonomy?

- Risk-awareness = ability to operate safely, robustly, and efficiently in a highly uncertain environment
- Capabilities that make autonomy risk-aware:
  - Explicitly consider ***uncertainty*** in the environment
  - Ability to ***balance*** risk-avoidance and performance
  - Ability to allow users to ***tune*** the balance
  - Talk: "Intro to Risk-aware Planning" by me at 2:10p
- Resilient Spacecraft Systems Project
  - FY15/16 Topic R&TD (PI: Mitch Ingham)
  - Developed architecture & algorithms for risk-aware autonomy
  - Focus: model-based planning based on stochastic optimization
  - Talk: "Resilient Spacecraft Executive" by Mitch Ingham at 2:30p

Machine Learning for Risk-aware Autonomy

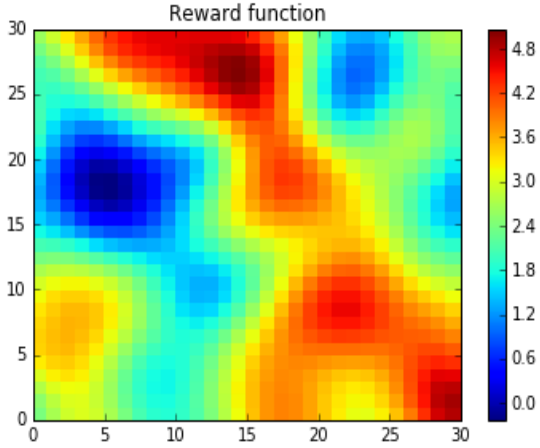
- Adapt to unknown/changing environment
  - How to build a model of safety of unknown environment?
  - How to predict safety and utility of unvisited destinations?
  - How to balance exploration and exploitation?
    - Exploration: discover unknowns
    - Exploitation: collect known rewards (e.g., science gain)
  - Approach: safe exploration (next slide)
- Quickly compute suboptimal but good-enough solution
  - Model-based stochastic optimization can find strictly optimal solution but requires intensive computation
  - Humans regularly perform risk-aware planning but do not solve stochastic optimization (at least explicitly)
    - Solutions are usually suboptimal but good enough
  - How to replicate human's ability of quickly making risk-aware decisions?
  - Approach: Imitation learning (two slides later)

Safe Exploration

Work by Akifumi Wachi and Yanan Sui

: Safe

: Uncertain/unsafe

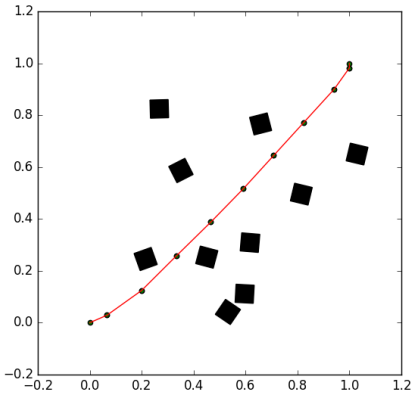


- An agent collects reward while exploring unknown world
- Each state of the world has:
  - Safety value, representing the level of safety. If this value is below a threshold, the state is unsafe
  - Utility value, representing the amount of reward that the agent can collect from the state
- Safety and utility values are unknown in the beginning
  - Except that the initial state is known to be safe
- Agent can predict the safety and utility values of nearby unvisited states
- Main trade-off: exploration vs exploitation
  - Too much exploration: keep exploring unknown states without collecting reward
  - Too much exploitation: Collect rewards only from known states, ignoring the possibility that more reward could be collected from unknown states
- Talks:
  - Yanan Sui (Caltech), 9:40a
  - Akifumi Wachi (U Tokyo), 10:00a
- Publication
  - "Safe Exploration and Optimization of Constrained MDPs using Gaussian Processes," A. Wachi, Y. Sui, Y. Yue, M. Ono. Submitted to IJCAI 2017

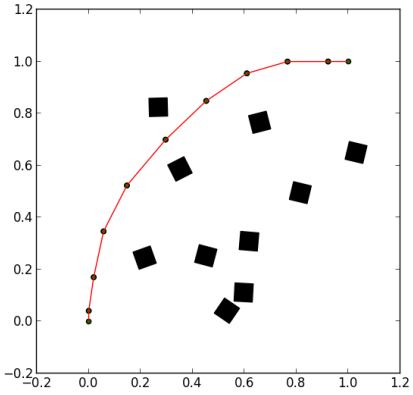
Imitation Learning w/ Monte-Carlo tree search

Work by Ravi Lanka Subrahmanya and Jialin Song

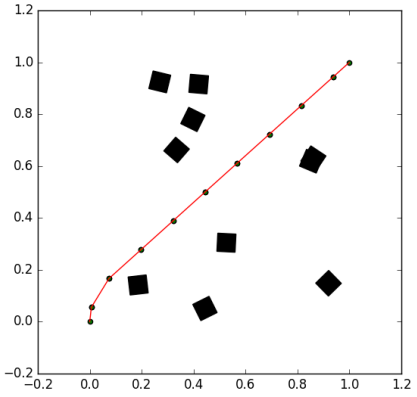
Optimal path



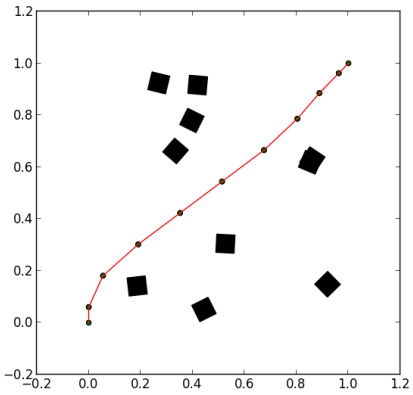
Imitation learning



Optimal path



Imitation learning



- Use model-based planner as a “teacher”
  - Used *pSulu* risk-aware path planner
  - Makes optimal decision by solving MILP (exponential computation time wrt # of time steps and obstacles)
  - Searches the optimal solution on a tree with branch-and-bound (BnB) method
- Agent learns the decisions of model-based planner by imitation learning
  - Intuition: make an intelligent guess about which branch in a tree is most promising
  - Uses Monte-Carlo tree search instead of BnB
  - Represented by a computationally efficient mode
- Talks
  - Jialin Song (Caltech), 10:30a
  - Ravi Lanka Subrahmanya (JPL), 10:50a
  - Hoang Le (Caltech), 11:10a
- Publication
  - Conference paper being prepared for submission in May, 2017

Workshop Agenda

Time	Speaker	Affiliation	Category	Talk title
9:00	Hiro Ono	JPL		Overview of Risk-aware Machine Learning Project
9:10	Yisong Yue	Caltech	Machine learning	Machine Learning Research at Caltech
9:40	Yanan Sui	Caltech	Machine learning	Exploration with safety and utility constraints
10:00	Akifumi Wachi	U Tokyo	Machine learning	Safe Exploration and Optimization of MDPs using Gausian Processes
10:20	Break			
10:30	Jialin Song	Caltech	Machine learning	ML for risk-aware path planning via Monte Carlo tree search and imitation learning
10:50	Ravi Lanka Subrahmanya	JPL	Machine learning	(TBD) ML-based solution for risk-aware path planning (2) - BnB approach
11:10	Hoang Le	Caltech	Machine learning	Smooth Imitation Learning
11:30	Lunch			
1:00	Shreyansh Daftry	JPL	Machine learning	Introspection: Risk-aware Perception for Autonomous Systems
1:20	Eric Timmons	MIT	Application	Risk-aware autonomy for AUVs
1:40	Akifumi Wachi	U Tokyo	Application	Automation Challenges for the Small Satellites of Univ. of Tokyo
2:00	Break			
2:10	Hiro Ono	JPL	Risk-aware planning	Intro to risk-aware planning
2:30	Mitch Ingham	JPL	Risk-aware planning	Resilient Spacecraft Executive
2:50	Raymond Francis	JPL	Application	AEGIS autonomous targeting for ChemCam: Protecting the rover and the instrument
3:10	Break			
3:20	Discussion: next big ideas			
5:00	Adjourn			